云南天文台手绘太阳黑子数据系统*

祝高飞[1,2]，郑胜[1]，林钢华[2]，曾祥云[1,2]，冯永利[3]，陶金萍[3]，舒瑶[1]
（1.三峡大学理学院，湖北 宜昌 443002；
2.中国科学院国家天文台太阳活动重点实验室，北京 100012；
3.中国科学院云南天文台，云南 昆明 650011）

摘 要：中国科学院云南天文台多年来积累了近16 000张手绘太阳黑子观测数据，建立完善的查询与统计系统对这些海量数据进行科学的管理与统计分析成为必要。数字化数据包括太阳黑子记录近9万个，有效记录数据100万余个。系统提供了一个数据管理、数据检索、数据统计分析的平台。借助数据管理系统，可对太阳黑子相关参数进行长周期的统计。

关键词：太阳黑子；数据管理；黑子信息检索；黑子统计分析

中图分类号：P182.2；TP311　　文献标识码：A　文章编号：1672-7673

各种太阳活动中，最为醒目也最容易观测的现象是太阳黑子。1610 年，国外开始用望远镜断断续续地对黑子进行观测。自1818 年开始才有比较常规的每日黑子观测，从而有比较可靠的黑子数据[1]。早期对太阳黑子观测的数据通常是手绘在纸张上，导致目前存在大量的纸质太阳黑子记录。由于纸质的易碎性,信息不便于检索，阻碍了天文台之间进行数据的整合和共享，因此需要将纸质图像转换为数字信息存储，以便于分析与查看。

我国太阳物理观测资料在具有时区优势，这使得我国太阳观测资料在国际上具有稀缺性。云南天文台手绘太阳黑子图像记录信息丰富，在长周期太阳活动相关研究中是不可多得的历史数据。云南天文台观测的1981-1992年黑子面积数据○，已被多次用于全球黑子数据的补充完善[Error: Reference source not found]。由于我国的手绘太阳黑子观测资料有其独特的价值，2014年由国家天文台怀柔基地牵头，对云南天文台和紫金山天文台近一个世纪的历史手绘黑子观测资料进行了系统的扫描、整理和数字化图像存储。为了更好地发挥历史观测资料的科学价值，方便地用于太阳活动的长期变化规律研究，需将扫描数字图像中的黑子信息准确可靠地提取并实现网络共享。利用图像分析技术对手写字符进行分割，以深度学习方法[5]进行字符识别从而实现手写黑子信息的自动提取[8]，实现太阳黑子手绘图的数字化。

将原始太阳黑子数字化数据进行系统的管理，从而能够通过数据库检索，查询有关黑子的信息，并且可以统计分析黑子各种变化情况，快捷方便，易于使用。最重要的是能够汇聚各个天文台的观测数据而进行有效的数据整理以便全面深入地研究太阳黑子的活动及其对地球的影响。因此，手绘太阳黑子图像数字化之后数据的系统化管理与分析显得尤为重要。本文利用MVC模式加上EF(Entity Framework)实体框架技术构建中国科学院云南天文台手绘太阳黑子数据系统，对数据进行集成化管理，并能对数据进行清晰、直观的统计分析。

1 数据获取

中国科学院云南天文台1958年至2015年的手绘太阳黑子图数字化扫描已完成。其中，每张手绘太阳黑子图像信息量1大，信息类型丰富多样，主要包括手绘太阳黑子图像绘制常规记录以及黑子信息，如图1。

*基金项目：国家自然科学天文联合基金 (U1531247)；科技部基础工作专项重点项目(2014FY120300,2012FY120500)；国家自然科学天文联合基金(U1331113)、中国科学院太阳活动重点实验室开放课题(KLSA201617)资助.
收稿日期：2017-06-12 ； 修订日期：2017-07-11
作者简介：祝高飞，男，硕士研究生，研究方向:图像处理，智能系统 Email:zhugaofei@nao.cas.com

○https://www.ngdc.noaa.gov/stp/solar/sunspotregionsdata.html

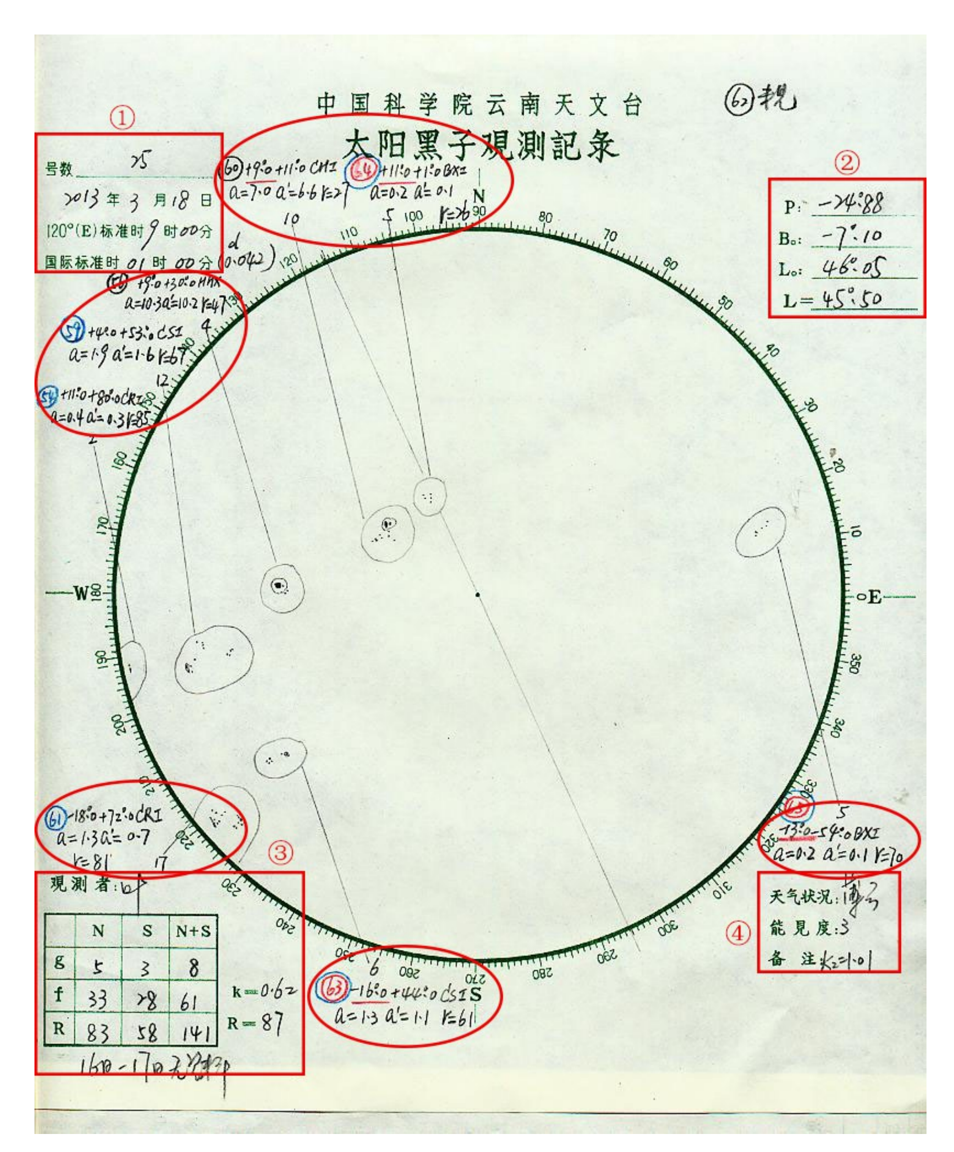


图1 扫描太阳黑子信息记录图像

Fig.1 Scan sunspot information to record images

矩形框1包含：号数（规定1853年11月9日本初子午圈转到日面中心的时刻为太阳的第1个自转周开始，自转周连续编号。每年中各个自转周的号数和开始日期都可从天文年历中查到）、观测日期、北京时间（东经120°标准时）、国际标准时（UTC）。

矩形框2包含：P 角（为自日面北点起的太阳自转轴北极的方位角）、B_0 和 L_0（分别为观测日世界标准时零时的日面中心纬度和经度、L（代表观测时刻的日面中心经度）。

矩形框3包含：北半球日面黑子群个数 gN、南半球日面黑子群个数 R、全日面黑子群个数 gNS、北半球日面黑子个数 fN、南半球日面黑子个数 fS、全日面黑子个数 fNS、北半球沃夫数 RN、南半球沃夫数 RS、全日面沃夫数 RNS、天文台因素 k、沃夫数 R。

矩形框4包含：天气状况、能见度、备注。

椭圆包含：黑子群编号、经度、纬度、黑子群结构类型、单个黑子群总面积、单个黑子群中最大黑子的面积(日面面积的百万分之一，简称最大黑子的面积)、黑子群质心到太阳投影圆中心的直线距离（毫米，简称半径）。

通过计算机辅助人工处理方式将扫描手绘太阳黑子图像中上述信息进行数字化处理，其中，文[11]详细介绍了深度学习方法在手绘太阳黑子图像中信息数字化中的应用，通过图像分割的方式，将手绘太阳黑子图像中手写字符进行分割，然后使用深度学习方法进行

字符识别，然后将数字化信息分别存储于两个数据表格中，记录格式如表1、表2，其中包含黑子信息记录近9万个，有效数据近100余万。

表1 手绘图常规记录数字化信息记录表

Table 1 Regular Information record table

号数	日期	北京时间	世界时间	天数	P	B0	Lo	L	gN	gS	gNS	fN	fS	fNS	RN	RS	RNS	k	R	Visible	K2
22	2013-03-11	09:15	01:15	0.052	-23.74	-7.23	138.32	137.63	5	1	6	24	32	56	74	42	116	0.62	72	3	1.01
23	2013-03-14	09:10	01:10	0.049	-24.27	-7.19	98.77	98.13	7	1	8	53	2	55	123	12	135	0.62	84	3	1.01
24	2013-03-15	08:45	00:45	0.031	-24.44	-7.17	85.59	85.17	7	2	9	43	10	53	113	30	143	0.62	89	3	1.01
25	2013-03-18	09:00	01:00	0.042	-24.88	-7.1	46.05	45.5	5	3	8	33	28	61	83	58	141	0.62	87	3	1.01
26	2013-03-21	09:05	01:05	0.045	-25.27	-7.02	6.5	5.9	3	0	3	21	0	21	51	0	51	0.62	32	3	1.01
27	2013-03-22	09:15	01:15	0.052	-25.39	-6.98	353.31	352.62	4	1	5	13	4	17	53	14	67	0.62	42	3	1.01
28	2013-03-25	09:20	01:20	0.056	-25.69	-6.87	313.76	313.03	2	1	3	7	7	14	27	17	44	0.62	27	3	1.01
29	2013-03-29	10:00	02:00	0.083	-26	-6.7	261	259.9	2	3	5	8	18	26	28	48	76	0.62	47	3	1.01
30	2013-05-13	09:00	01:00	0.042	-21.47	-2.89	26.65	26.1	4	4	8	79	31	110	119	71	190	0.62	118	3	1.01

表2 手绘图黑子数字化信息记录表

Table 2 Sunspot information record table

日期	黑子号	经度	纬度	类型	黑子群	最大黑	半径	黑子数
20130318	54	11	80	CRI	0.4	0.3	85	2
20130318	56	9	30	HHX	10.3	10.2	47	4
20130318	59	4	53	CSI	1.9	1.6	67	12
20130318	60	9	11	CHI	7	6.6	27	10
20130318	61	-18	72	CRI	1.3	0.7	81	17
20130318	63	-16	44	CSI	1.3	1.1	61	6
20130318	64	11	1	BXI	0.2	0.1	26	5
20130318	65	-13	-54	BXI	0.2	0.1	70	5

2 系统设计与实现

2.1 功能设计

系统目前已经构建完成，不仅可提供太阳黑子观测数据检索下载，还可对系统中的数据进行多方面的统计和分析，研究人员可直观了解太阳黑子相关参数的变化趋势。系统主要分 3 个大功能模块以及多个小的细分模块，如图 2。

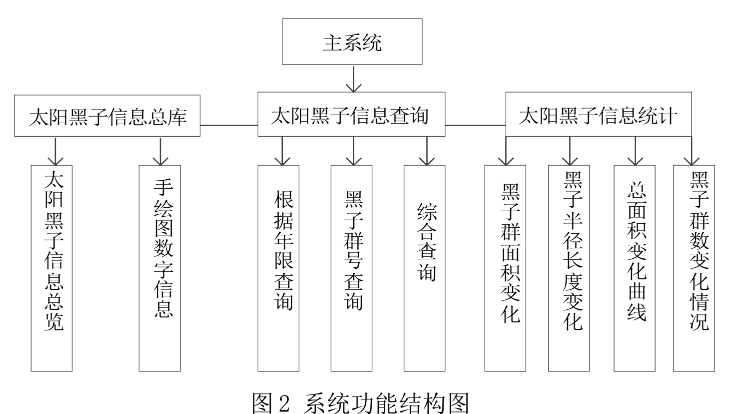


图 2 系统功能结构图

Fig.2 System function chart

2.2 架构设计

系统采用 MVC 的系统设计模式进行构建，MVC 是模型(model)－视图(view)－控制器(controller)的缩写，用一种业务逻辑、数据、界面显示分离的方法组织代码，将业务逻辑聚集到一个部件里面，在改进和个性化定制界面及用户交互的同时，不需要重新编写业务逻辑。系统结构如图 3。

2.2.1 模型类的设计

模型类的设计为了使系统拥有一个可以方便使用与访问的实体类，可以供控制器的操

作和视图的调用，系统中设计了两个主要的模型类以及多个辅助视图类，其中两个主要模型类：

（1）基本信息类

基本信息类的设计主要是映射手绘太阳黑子图中具有一对一关系的数据信息，比如日期、编号、标准时间、国际时间、天气状况、能见度、备注等的数据信息。

（2）黑子信息类

黑子信息类的设计主要是映射手绘太阳黑子图中每个黑子（群）的各项黑子参数信息，包括黑子的编号、单个黑子群总面积、最大黑子的面积、经度、纬度、半径等信息。

2.2.2 控制器的设计

控制器负责处理用户请求，然后调用相应的视图显示。系统包含 3 个主要的控制器以及多个辅助控制器，其中 3 个主控制器：

（1）太阳黑子总控制器

太阳黑子总控制器主要控制太阳黑子总库中的手绘太阳黑子图的各种信息的展示，包括手绘图的各种信息列表显示、太阳黑子信息的详细显示等。

（2）太阳黑子查询控制器

太阳黑子查询控制器负责根据用户请求分别处理不同参数的数据检索任务，例如用户根据年份、黑子群号、单个黑子群总面积、最大黑子的面积、经度、纬度等各种信息检索相应的太阳黑子信息，通过相应视图反馈给用户。

（3）太阳黑子统计控制器

太阳黑子统计控制器相当于处理分析控制器，主要负责统计分析太阳黑子的单个黑子群总面积、最大黑子的面积、半径、以及黑子群总面积和数量的变化情况等，以图表的方式直观的展示给用户。

2.2.3 视图的设计

视图的设计也可以称作是 UI 界面设计，主要是负责接收来自控制器传来的数据，以一种友好的方式展示给用户使用。系统包括主视图、共享视图、部分视图以及各种相应控制器的展示视图。

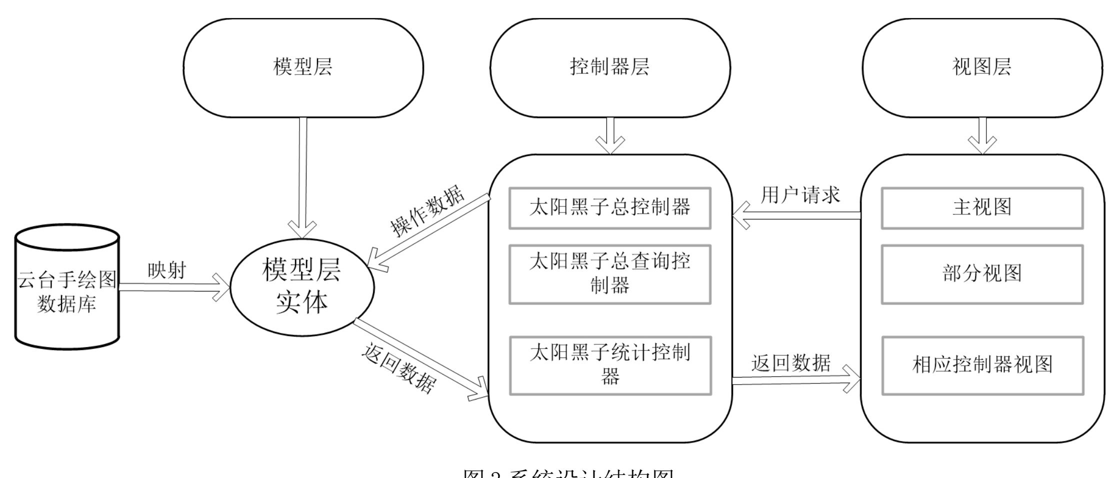


图 3 系统设计结构图

Fig.3 System design

3 系统功能

3.1 太阳黑子信息总库

太阳黑子信息总库是将云南天文台所有手绘太阳黑子图的数字化信息以年份为导航，

将每年所有观测的手绘太阳黑子图以列表的形式展示给用户。图 4 为 1991 年数据展示界面。查询结果中双击每张手绘图将展示该图及相应的数字化信息如图 5，并且通过鼠标交互可以将局部图像显示在右侧显示框中。



图 4　1991 年太阳黑子信息列表

Fig. 4 1991 sun sunspot information list



图 5　1991 年 1 月 1 日太阳黑子信息显示及局部视图

Fig. 5 January 1, 1991 sun sunspot information display and local view

3.2 太阳黑子信息查询

太阳黑子信息查询模块包含：年限查询、黑子群号查询、综合查询，以日期、黑子群号、经度、纬度、单个黑子群总面积、最大黑子的面积等，经过服务器中的太阳黑子信息查询控制器处理之后返回符合筛选条件的太阳黑子信息。以综合查询为例，选择日期为 1990 年 1 月 1 日到 2000 年 1 月 1 日，单个黑子群总面积范围 0.5 到 1.5 为筛选条件，处理得到的结果如图 6。



日期	黑子群号	经度	纬度	类型	单个黑子群总面积	最大黑子面积	半径	黑子数
1991-01-01	1	1	38.5	DRO	0.9	0.5	54	7
1991-01-01	654	-21	68	DRO	1.5	1	81	3
1991-01-01	666	-24	-9	CRI	1.4	0.5	34	31
1991-01-02	4	-18	-75	DRO	1.1	0.4	84	3
1991-01-02	666	-24	3.5	CRI	1.1	0.3	31	26

图 6　综合查询的部分结果

Fig.6 Part of the results of the comprehensive query

3.3 太阳黑子信息统计

太阳黑子信息统计模块又分为：黑子群面积（包含单个黑子群总面积和最大黑子的面积）变化曲线、黑子群半径长度变化曲线、黑子群总面积变化情况、黑子数变化情况，可对年度观测数据进行处理，也可对几年甚至百年数据进行统计分析。以 1991 年全年黑子信息为例，经系统处理后，黑子总面积变化情况如图 7，黑子个数统计情况如图 8。

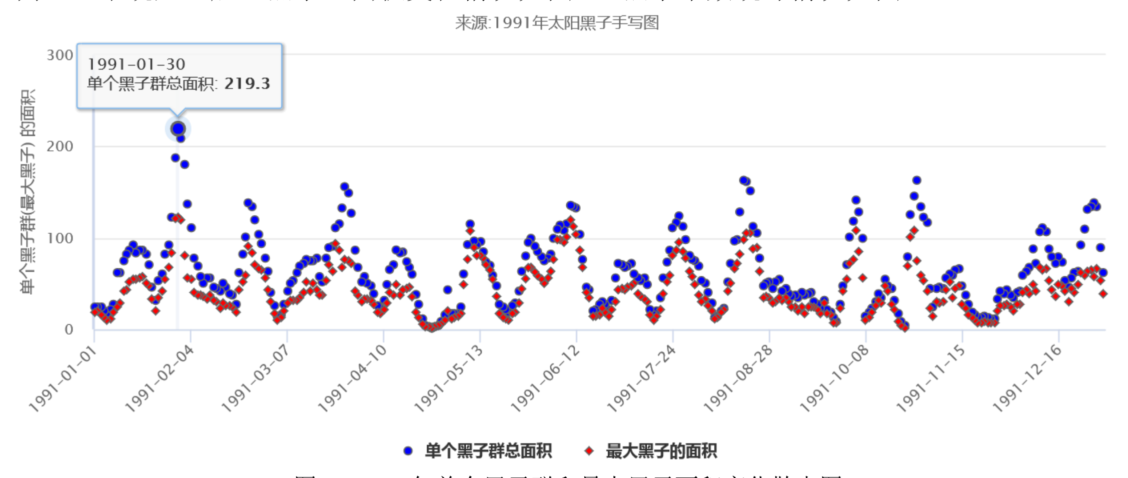


图 7　1991 年单个黑子群和最大黑子面积变化散点图

Fig.7 1991 sunspot group area change scatter plot

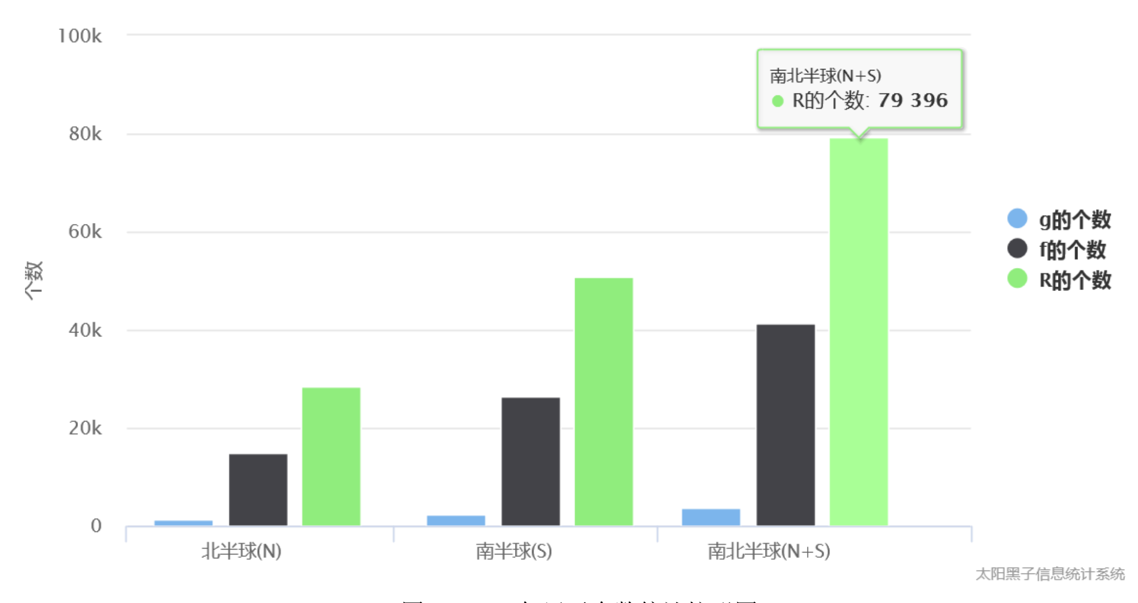


图 8　1991 年黑子个数统计柱形图

Fig.8 The number of sunspots in 1991

3.4 长周期黑子信息统计分析

目前，系统中已经导入 1993 年至 2015 年的手绘太阳黑子信息，借助系统，对这些数据进行统计，研究人员可清晰直观看到每年太阳黑子相关物理参数的变化情况。通过系统对 1993 年-2015 年的太阳黑子相关属性值进行统计，结果如表 3。

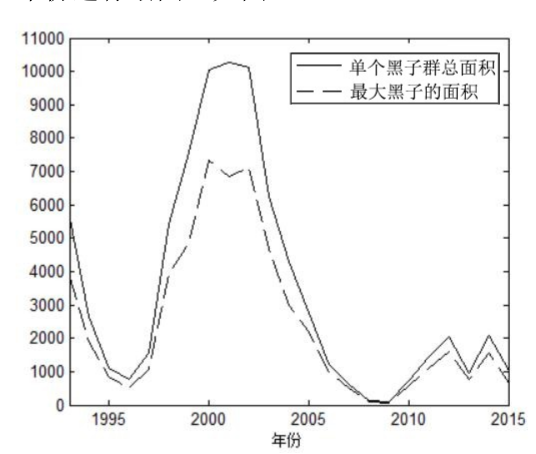
表 3　1993 年至 2015 年日面太阳黑子相关参数统计

Table 3　From 1993 to 2015 sun sunspot related parameters statistics

年份	单个黑子群总面积	最大黑子的面积	黑子群个数	黑子个数
1993	5592.5	3817.3	1560	12890
1994	2642.9	1884.6	1024	6913
1995	1100.2	861.6	555	3706
1996	785.1	510.5	297	1924
1997	1573.4	1056.1	605	4559
1998	5418.9	3933.8	1652	12556
1999	7565.7	4869.8	2428	21718
2000	10049.5	7325.9	3056	27644
2001	10254.4	6866.9	3811	41286
2002	10113.6	7124.6	2939	24774
2003	6261	4676	1814	13907
2004	4290.9	3001.1	1131	9277
2005	2751.4	2214.2	745	5866
2006	1210.7	1007.1	361	2870
2007	621.7	511	210	1642
2008	135.3	95.6	91	593
2009	99.1	62.8	51	373
2010	726.3	572	328	2095
2011	1458	1102.5	363	3096

2012	2044.6	1598.7	634	5320
2013	975.6	766	343	2949
2014	2070.3	1557.1	504	4509
2015	1039.7	654.3	388	4692

通过数据可视化的方式，将表 3 中的单个黑子群总面积、最大黑子的面积、个数进行按年份进行绘图，如图 9。



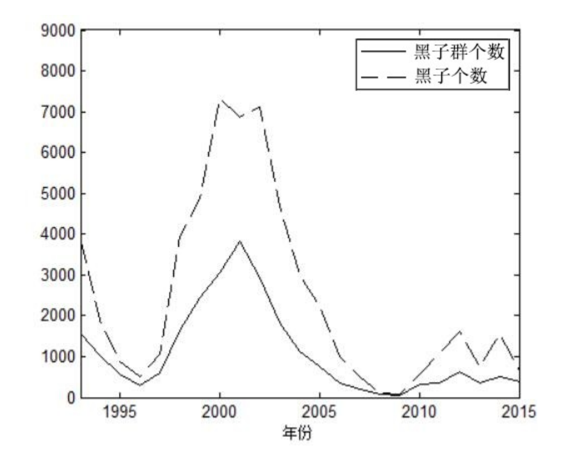


图 9　1993-2015 年黑子群相关属性图示

由图 9 可以看出，系统对于黑子面积与黑子（群）个数的统计，具有明显的波峰波谷特征，黑子数与黑子群数的多少，与黑子面积的变化趋势相吻合，黑子面积的大小及黑子数目的多少直接联合表征着太阳活动的剧烈程度。这种明显的特征现象与太阳活动周期变化规律相一致，对于研究太阳活动具有重要意义，同时说明系统中存储的数据具有一定的科学性和可靠性。

4 结论

云南天文台手绘太阳黑子数据系统目前已经完成手绘太阳黑子图总库、黑子信息查询、黑子信息统计分析等功能的建设。方便科研工作者按照不同的条件对黑子信息进行查询，借助系统可对太阳黑子的相关活动进行可视化研究，系统功能可进一步扩展，能给太阳物理科研人员提供便捷的服务，对于推动太阳黑子活动相关研究具有重要意义。后期将着手拓展太阳黑子运动轨迹绘制、太阳黑子图像三维展示等系统功能，能更加方便、直观、清晰地应用于长周期太阳黑子活动的研究。

致谢

感谢中国科学院云南天文台为本文实验提供手绘太阳黑子图像数据，感谢相关工作人员对工作的建议和支持。

参考文献:

[1] 李可军, 苏同卫, 梁红飞. 现代黑子观测的太阳黑子活动的周期性[J]. 科学通报, 2004, 49(24):2511-2516.
Li Kejun, Su Tongwei, Liang Hongfei. Periodicity of sunspot activity observed by modern sunspots [J]. Chinese Science Bulletin, 2004, 49 (24):

2511-2516.

[2] Baranyi T, Gyori L, Ludmány A, et al. Comparison of sunspot area data bases[J]. Monthly Notices of the Royal Astronomical Society, 2001, 323(1):223-230.

[3] Győri L. Study of differences between sunspot and white light facular area data determined from SDO/HMI and SOHO/MDI observations[J]. Solar Physics, 2012, 280(2):365-378.

[4] Balmaceda L A, Solanki S K, Krivova N A, et al. A homogeneous database of sunspot areas covering more than 130 years[J]. Journal of Geophysical Research Atmospheres, 2009, 114(A7):431-433.

[5] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks[J]. Journal of Machine Learning Research, 2011, 15：315-323.

[6] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.

[7] Yu D, Deng L. Deep learning and its applications to signal and information processing [Exploratory DSP][J]. IEEE Signal Processing Magazine, 2011, 28(1):145-154.

[8] Niu X X, Suen C Y. A novel hybrid CNN-SVM classifier for recognizing handwritten digits[J]. Pattern Recognition, 2012, 45(4):1318-1325.

[9] Zheng S, Zeng X Y, Lin G H, et al. Sunspot drawings handwritten character recognition method based on deep learning[J]. New Astronomy, 2016, 45:54-59.

Yunnan Observatories Hand-painted Sunspot Data System

Zhu Gaofei[1,2], Zheng Sheng[1], Lin Ganghua[2], Zeng Xiangyun[1,2], Feng Yongli[3], Tao Jinping[3], Shu Yao[1]

(1. College of Science, China Three Gorges University, Yichang 443002, China；

2. Key Laboratory of Solar Activity, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China;3.Yunnan Observatories,Chinese Academy of Sciences,Kunming 650011,China)

Abstract: Over the years, the Yunnan Observatory of Chinese Academy of Sciences has accumulated nearly 16,000 sunspot drawings. It is necessary to establish a complete system of query and statistics for scientific management and statistical analysis of these massive data. Digital data, including sunspots recorded nearly 90,000, the effective record of more than 100 million data. The system provides a platform for data management, data retrieval, and statistical analysis of data. With the data management system, the sunspot related parameters can be long-term statistics.

Key words: Sunspot;Data management;Sunspot information retrieval;Sunspot statistical analysis;